APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/614,111	07/03/2003	Daniel Dulitz	60963-0005-US	7663

24341	7590	08/24/2006

MORGAN, LEWIS & BOCKIUS, LLP.
2 PALO ALTO SQUARE
3000 EL CAMINO REAL
PALO ALTO, CA 94306

EXAMINER
MORRISON, JAY A

ART UNIT	PAPER NUMBER
2168	

DATE MAILED: 08/24/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

	Application No.	Applicant(s)
Office Action Summary	10/614,111	DULITZ ET AL.
	Examiner	Art Unit	
	Jay A. Morrison	2168	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE <u>3</u> MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.
- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☐ Responsive to communication(s) filed on *15 June 2006*.

2a)☒ This action is **FINAL**.   2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *12-20,37-40 and 42-58* is/are pending in the application.

    4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *12-20,37-40 and 42-58* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *03 July 2003* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

    Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

    Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

    a)☐ All b)☐ Some * c)☐ None of:

      1.☐ Certified copies of the priority documents have been received.

      2.☐ Certified copies of the priority documents have been received in Application No. _____.

      3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

    * See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1)☒ Notice of References Cited (PTO-892)
2)☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3)☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.

4)☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5)☐ Notice of Informal Patent Application (PTO-152)
6)☐ Other: _____.

DETAILED ACTION

1.   Claims 12-20 and 37-58 are pending.

Claim Rejections - 35 USC § 102

2.   The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that

form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

3.   Claims 12-17,40-48 and 50-55 rejected under 35 U.S.C. 102(e) as being

anticipated by <u>Meyerzon et al.</u> ('<u>Meyerzon</u>' hereinafter) (Patent Number 6,547,829).

As per claim 12, <u>Meyerzon</u> teaches

"constructing a plurality of tables, each table corresponding to a portion of a

document address space " (builds new index based on documents, column 4, lines 43-

60) ", storing information identifying documents having a same document identifier and

each identified document having an associated document rank" (column 2, lines 3-16);

"receiving a newly crawled document, such document characterized by a

document identifier and a document rank" (column 2, lines 3-16);

"reading information stored in the plurality of tables to identify a set of documents, if any, sharing the document identifier of the newly crawled document" (column 9, lines 18-29);

"updating the information stored in at least one of the tables in accordance with the document ranks of the identified set of documents and the newly crawled document" (column 2, lines 3-16);

"and determining a representative document for the newly crawled document and the identified set of documents" (column 9, lines 32-40).

As per claim 13, Meyerzon teaches

"information identifying the identified set of documents, including a particular document serving as a representative document of the identified set, is stored in one or more tables" (column 9, lines 32-40).

As per claim 14, Meyerzon teaches

"comparing the document rank of the newly crawled document with that of the particular document from the identified set in accordance with a set of predefined comparison criteria; selecting the newly crawled document as the representative document if the set of predefined comparison criteria are met" (column 5, lines 20-40);

"and keeping the particular document as the representative document if the set of predefined comparison criteria is not met" (column 2, lines 32-40).

As per claim 15, <u>Meyerzon</u> teaches

"the set of predefined comparison criteria comprise at least two parameters, one parameter for comparison with an absolute difference of document ranks between the newly crawled and the particular document, and another parameter for comparison with a ratio of document ranks between the newly crawled document and the particular document" (column 5, lines 20-40).

As per claim 16, <u>Meyerzon</u> teaches

"the updating includes inserting information identifying the newly crawled document into the at least one table only when a predefined insertion condition is satisfied" (column 9, lines 32-40).

As per claim 17, <u>Meyerzon</u> teaches

"the predefined insertion condition is that the document rank of the newly crawled document is higher than the document rank of at least on document in the identified set of documents" (column 2, lines 32-40).

As per claim 40, <u>Meyerzon</u> teaches

"constructing a plurality of data structures for storing information of documents" (builds new index based on documents, column 4, lines 43-60) ", each document characterized by a document identifier and a document rank, the information stored in the plurality of data structures include the document identifier and a document rank for

each document" (URL in history table and CID in separate CID table, column 2, lines 64

through column 3, line 22);

"receiving a requesting document in association with its document identifier and

document rank" (column 2, lines 3-16);

"selecting from the plurality of data structures a set of documents, if any, sharing

the same document identifier as the requesting document" (column 9, lines 18-40);

"generating a new set of documents from the requesting document and the

selected set of documents in accordance with their document rank" (column 2, lines 3-

16);

"identifying a representative document of the new set of documents" (column 9

lines 32-40).

As per claim 41, Meyerzon teaches

"the score information for each document includes a document rank metric"

(column 2, lines 3-16).

As per claim 42, Meyerzon teaches

"the plurality of data structures include a data structure for storing information of

multiple sets of documents, each set of documents sharing a same document content"

(column 2, line 64 through column 3, line 22).

As per claim 43, Meyerzon teaches

"the plurality of data structures include a data structure for storing information of

multiple sets of documents, each set of documents sharing a same document address"

(storage location, column 2, line 64 through column 3, line 22).

As per claim 44, <u>Meyerzon</u> teaches

"the document identifier is a fixed length fingerprint of document content of a

document characterized by the document identifier" (content identifier, column 2, line 64

through column 3, line 22).

As per claim 45, <u>Meyerzon</u> teaches

"the document identifier is a fixed length fingerprint of an address of a document

characterized by the document identifier" (content identifier, column 2, line 64 through

column 3, line 22).

As per claim 46, <u>Meyerzon</u> teaches

"sorting the requesting document and the selected set of documents in

accordance with a metric included in the score information of the requesting document

and selected set of documents; and selecting a new set of documents, having at most a

predefined number of documents from the requesting document and the selected set of

documents based on the sorting result" (column 2, lines 3-16).

As per claim 47, <u>Meyerzon</u> teaches

"the score information for each document includes a document rank" (column 2, lines 3-16);

"comparing the document rank of the requesting document with that of a particular document from the selected set of documents in accordance with a set of predefined comparison criteria, wherein the particular document was previously determined to be the representative document for the selected set of documents" (column 5, lines 20-40);

"selecting the requesting document as the representative document for the new set of documents if the set of predefined comparison criteria are met" (column 2, lines 32-40);

"and keeping the particular document as the representative document for the new set of documents if the set of predefined comparison criteria is not met" (column 2, lines 32-40).

As per claim 48, Meyerzon teaches

"the set of predefined comparison criteria comprise at least two parameters, one parameter for comparison with an absolute difference of document rank between the requesting document and the particular document, and another parameter for comparison with a ratio of document rank between the requesting document and the particular document" (column 8, lines 39-61).

As per claims 50-55,

These claims are rejected on grounds corresponding to the arguments given

above for rejected claims 12-17 and are similarly rejected.

Claim Rejections - 35 USC § 103

4.   The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all

obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set
> forth in section 102 of this title, if the differences between the subject matter sought to be patented and
> the prior art are such that the subject matter as a whole would have been obvious at the time the
> invention was made to a person having ordinary skill in the art to which said subject matter pertains.
> Patentability shall not be negatived by the manner in which the invention was made.

This application currently names joint inventors. In considering patentability of

the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of

the various claims was commonly owned at the time any inventions covered therein

were made absent any evidence to the contrary. Applicant is advised of the obligation

under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was

not commonly owned at the time a later invention was made in order for the examiner to

consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g)

prior art under 35 U.S.C. 103(a).

5.   Claims 18-20, 37-39 and 56-58 are rejected under 35 U.S.C. 103(a) as being

unpatentable over Meyerzon et al. ('Meyerzon' hereinafter) (Patent Number 6,547,829)

in view of Rujan et al. ('Rujan' hereinafter) (Patent Number 6,976,207).

As per claim 18, Meyerzon teaches

"constructing a plurality of tables, each table corresponding to a segment of a document address space, storing information identifying documents having a same document identifier and each identified document having an associated document rank, wherein the plurality of tables comprise N+1 tables where N is an integer greater than one, wherein the N+1 tables comprise N tables, each generated during a respective phase of a set of N crawling phases, and a current table generated during a current one of the N crawling phases wherein an oldest one of the N tables was generated during a previous instance of the current crawling phase" (column 4, lines 43-60);

"receiving a newly crawled document, such document characterized by a document identifier and a document rank" (column 2, lines 3-16);

"reading information stored in the N+1 tables to identify a set of documents, if any, sharing the document identifier of the newly crawled document" (column 4, lines 43-60);

"updating the information stored in the current table in accordance with the document rankings of the identified set of documents and the newly crawled document" (column 4, line 43 through column 5, line 13);

"determining a representative document for the newly crawled document and the identified set of documents" (column 2, lines 32-40);

"and upon completion of the current crawling phase, ... of the N tables" (column 5, lines 1-20).

Meyerzon does not explicitly indicate "retiring the oldest one".

However, Rujan discloses "retiring the oldest one" (column 15, lines 20-25).

It would have been obvious to one of ordinary skill in the art to combine

Meyerzon and Rujan because using the steps of "retiring the oldest one" would have

given those skilled in the art the tools to create an effective information storage and

retrieval system. This gives the user the advantage of keeping a limited amount of

historic information.

As per claim 19, Meyerzon teaches

"the reading comprises reading from a merged table that stores information from

a plurality of the N tables, and reading from the current table" (column 4, lines 43-60).

As per claim 20, Meyerzon teaches

"information identifying the identified set of documents, including a particular

document serving as a representative document of the identified set, is stored in one or

more tables" (column 9 lines 32-40).

As per claims 37-39,

These claims are rejected on grounds corresponding to the arguments given

above for rejected claims 18-20 and are similarly rejected.

As per claims 56-58,

These claims are rejected on grounds corresponding to the arguments given

above for rejected claims 18-20 and are similarly rejected.

6.   Claim 49 is rejected under 35 U.S.C. 103(a) as being unpatentable over

Meyerzon et al. ('Meyerzon' hereinafter) (Patent Number 6,547,829) in view of Lambert

et al. ('Lambert' hereinafter) (Publication Number 2002/0038350).

        As per claim 49,

        Meyerzon does not explicitly indicate "a document is a temporary redirect page

comprising a document content, a source document address, and a target document

address".

        However, Lambert discloses "a document is a temporary redirect page

comprising a document content, a source document address, and a target document

address" (paragraph [0057]).

        It would have been obvious to one of ordinary skill in the art to combine

Meyerzon and Lambert because using the steps of "a document is a temporary redirect

page comprising a document content, a source document address, and a target

document address" would have given those skilled in the art the tools to accurately

represent web sites and the content that they hold. This gives the user the advantage of

recognizing web page structure.

                            Response to Arguments

        Applicant's arguments filed 6/15/06 have been fully considered but they are not

persuasive.

With regards to Applicant's argument that <u>Meyerzon</u> does not teach or anticipate

claim 12 because <u>Meyerzon</u> does not use document rank for use in detecting or

processing duplicate documents, it is respectfully noted that the instant claim is directed

to "storing information identifying documents having a same document identifier" and

"an associated document rank", "reading information stored in the plurality of tables to

identify a set of documents, if any, sharing the document identifier of the newly crawled

document" and then "updating the information stored in at least one of the tables in

accordance with the document ranks of the identified set of documents and the newly

crawled documents". The claims do not explicitly state that the document rank is used in

detecting or processing duplicate documents; it seems that the document identifier is

more instrumental in that process. Granted, the document rank is used for "updating the

information stored in at least one of the tables", but the document identifier is used for

reading the documents in the tables which match the newly crawled document.

Meyerzon does teach that the document identifier is used in detecting or

processing duplicate documents (column 9, lines 18-29). It is therefore respectfully

noted that <u>Meyerzon</u> does in fact teach the claim.

With regards to Applicant's argument that <u>Meyerzon</u> does not teach or anticipate

claim12 because <u>Meyerzon</u> keeps no record of any newly acquired documents thought

to be duplicates and does not read information in tables to identify a set of duplicate

documents, it is noted that Meyerzon does in fact disclose keeping record of new

acquired documents thought to be duplicates (URL and CID committed to history table, column 9, lines 45-50) and reads information into tables to identify a set of duplicate documents (CID determined and URL and CID committed to history table, column 9, lines 32-50).

With regards to Applicant's argument that Meyerzon does not teach or anticipate claim 14 because Meyerzon comparison of two duplicate documents, it is noted that Meyerzon discloses that CID values are compared (column 9, lines 32-39) which discloses the limitation.

With regards to Applicant's argument that Meyerzon in view of Rujan does not teach or anticipate claims 18,27 and 36 because Meyerzon does not store information identifying documents having the same identifier and does not allow for the selection of a representative document from among the newly crawled document and the identified set of documents, it is noted that Meyerzon does in fact stored documents having the same identifier (URL in history table) and allows for the selection of a representative document (column 9, lines 32-50).

Applicant should submit an argument under the heading "Remarks" pointing out disagreements with the examiner's contentions. Applicant must also discuss the references applied against the claims, explaining how the claims avoid the references or distinguish from them.

With regards to Applicant's argument that <u>Meyerzon</u> in view of <u>Lambert</u> does not

teach or anticipate claims 18, 27 and 36 because <u>Lambert</u> does not teach storing

information identifying a set of documents having the same document identifier,

updating such information based on document rankings, or selecting a representative

document from among a newly crawled document and an identified set of documents, it

is respectfully submitted that the Applicant has not argued these claims.

Conclusion

1.   **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time

policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE

MONTHS from the mailing date of this action. In the event a first reply is filed within

TWO MONTHS of the mailing date of this final action and the advisory action is not

mailed until after the end of the THREE-MONTH shortened statutory period, then the

shortened statutory period will expire on the date the advisory action is mailed, and any

extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of

the advisory action. In no event, however, will the statutory period for reply expire later

than SIX MONTHS from the mailing date of this final action.

The prior art made of record, listed on form PTO-892, and not relied upon is

considered pertinent to applicant's disclosure.

Any inquiry concerning this communication or earlier communications from the

examiner should be directed to Jay A. Morrison whose telephone number is (571) 272-

7112. The examiner can normally be reached on M-F 8-4:30.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's

supervisor, Tim Vo can be reached on (571) 272-3642. The fax phone number for the

organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the

Patent Application Information Retrieval (PAIR) system. Status information for

published applications may be obtained from either Private PAIR or Public PAIR.

Status information for unpublished applications is available through Private PAIR only.

For more information about the PAIR system, see http://pair-direct.uspto.gov. Should

you have questions on access to the Private PAIR system, contact the Electronic

Business Center (EBC) at 866-217-9197 (toll-free).

TIM VO
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2100

Jay Morrison   Tim Vo
TC2100   TC2100